

KHUYNH HƯỚNG NGHIÊN CỨU NHÂN VĂN SỐ TỪ GÓC ĐỘ PHÂN TÍCH TRẮC LƯỢNG THƯ MỤC

TS Ninh Thị Kim Thoa

Trường Đại học KHXH&NV, Đại học Quốc gia Tp. Hồ Chí Minh

Tóm tắt: Bài viết trình bày kết quả nghiên cứu sử dụng phương pháp trắc lượng thư mục để phân tích một cách hệ thống các chủ đề và từ khóa của các tài liệu về nhân văn số. Kết quả nghiên cứu cũng cho thấy tình hình hợp tác nghiên cứu về nhân văn số giữa các tác giả và các quốc gia. Mặc dù các tác giả từ các nước đang phát triển đã có nhiều công trình nghiên cứu liên quan, các tác giả đến từ các quốc gia phát triển vẫn đóng vai trò nổi bật nhất trong nghiên cứu về nhân văn số. Kết quả nghiên cứu được trình bày trong bài viết này phần nào giúp các nhà khoa học có sự hiểu biết tổng quát về tình hình nghiên cứu nhân văn số hiện nay, cũng như tìm được những chủ đề nghiên cứu phù hợp.

Từ khóa: Nhân văn số; trắc lượng thư mục; VOSViewer, Scopus.

TRENDS IN DIGITAL HUMANITIES RESEARCH FROM THE PERSPECTIVE OF BIBLIOMETRIC ANALYSIS

Abstract: This article presents the research results using bibliometric method to systematically analyze topics and keywords of documents on digital humanities. The research results also show the situation of research collaboration among individual authors and countries. Although authors from developing countries have had many have relevant studies in the field, authors from developed countries still play the most prominent role in research on the digital humanities. The research results presented in this article partly help scientists have a general understanding of the current situation of digital humanities research, as well as find suitable research topics.

Keywords: Digital humanities; bibliometrics; VOSViewer, Scopus..

Đặt vấn đề

Thế giới ngày nay đang vận hành trong thời đại kỹ thuật số, với sự ứng dụng của các nền tảng công nghệ đột phá vào trong mọi lĩnh vực cuộc sống. Mỗi một công nghệ hay ứng dụng mới ra đời đều có thể giúp tăng khả năng của con người, gây ảnh hưởng đến xã hội, đến các hành vi của con người và xã hội. Điều này đặt ra những yêu cầu mới cho việc nghiên cứu và giảng dạy các lĩnh vực khoa học xã hội và nhân văn (KHXHNV), những ngành khoa học nghiên cứu về con người, mối quan hệ giữa con người với con người, giữa con người với xã hội. Nếu trước kia các ngành khoa học này thường sử dụng các phương pháp nghiên cứu truyền thống để thu thập, tổ chức, xử lý và trình bày dữ liệu với các thông tin, hiện vật và minh chứng thu thập được từ thực

tế thì ngày nay, việc nghiên cứu còn cần được thực hiện trong môi trường kỹ thuật số, bằng cách dùng những phương pháp mới và cách tiếp cận khác trước. Từ đó, một lĩnh vực nghiên cứu mới là nhân văn số (digital humanities) đã hình thành và phát triển. Sử dụng công cụ Google Scholar để tìm “mọi lúc” theo cú pháp tìm chính xác từ khóa “digital humanities” xuất hiện bất cứ nơi nào trong các bài viết đã cho 74.100 kết quả (ngày 03/3/2021), trong khi đó vào năm 2017 thì chỉ có 18.900 kết quả. Như vậy, trong khoảng 5 năm gần đây, số lượng các kết quả đã tăng gần gấp 4 lần. Xu hướng tìm kiếm thuật ngữ “digital humanities” trên toàn thế giới cũng có sự gia tăng mạnh mẽ từ khoảng cuối năm 2013 đến nay (xem Hình 1). Tuy nhiên, không tìm thấy dữ liệu thống kê tìm kiếm từ Việt Nam.



Hình 1. Xu hướng tìm kiếm thuật ngữ “digital humanities” trên Google trends

Mình chứng này cho thấy sự gia tăng nhanh chóng của các nghiên cứu liên quan đến nhân văn số cũng như sự quan tâm đến lĩnh vực này của cộng đồng các nhà khoa học trên thế giới, đồng thời cũng chứng minh hiện trạng tại Việt Nam. Vì vậy, việc khảo sát và phân tích một cách toàn diện các tài liệu liên quan đến lĩnh vực này nhằm cung cấp một bức tranh tổng quan, cũng như tìm ra những lĩnh vực hay chủ đề cụ thể cần được khai phá sâu hơn là một tiếp cận cần được quan tâm.

Bài viết trình bày một nghiên cứu được thực hiện dựa trên phương pháp nghiên cứu trắc lượng thư mục với mục đích cung cấp một phân tích tổng quan về chủ đề nhân văn số hiện nay.

1. Hiện trạng nghiên cứu về nhân văn số

Thuật ngữ “nhân văn số” bắt đầu từ khái niệm “điện toán nhân văn” (humanities computing), thể hiện việc sử dụng máy tính để phân tích dữ liệu nghiên cứu trong các ngành KHXXNV được bắt đầu từ năm 1949 với sự phát triển của các dự án đầu tiên như Index Thomisticus để ứng dụng điện toán nhân văn (humanities computing) trong ngôn ngữ học [Busa, 1980; Hockey, 2004]. Với sự phát triển của công nghệ kỹ thuật số, khái niệm nhân văn số xuất hiện, được hiểu là những hoạt động học thuật trong các lĩnh vực KHXXNV có áp dụng các công cụ kỹ thuật số và các phương pháp hỗ trợ của các phần mềm máy tính để nghiên cứu và giảng dạy [Gavin & Smith, 2012].

Nhân văn số chỉ mới thực sự thành một chủ đề nghiên cứu nổi bật trong một vài thập niên gần đây, đặc biệt cùng với sự ra đời của cuốn

sách “A companion to Digital Humanities” (Schreibman và cộng sự, 2004), đề cập đến lịch sử phát triển của nhân văn số, việc ứng dụng công nghệ điện toán trong nghiên cứu KHXXNV nói chung, và trong các lĩnh vực cụ thể như khảo cổ học, nghệ thuật, ngôn ngữ, âm nhạc...; các nguyên tắc và ứng dụng công nghệ trong nghiên cứu nhân văn số cũng như quá trình sản xuất, tạo lập, phổ biến và lưu trữ các nguồn lực số. Kể từ đó, các nghiên cứu về nhân văn số đã phát triển thành các hướng rộng hơn và phức tạp hơn, hiện đang được xem là một hình thức nghiên cứu nhân văn liên ngành, được phát triển để tăng cường và chuyển đổi học thuật nhân văn truyền thống thông qua các phương tiện kỹ thuật số.

Ngày nay, nhân văn số đã trở thành chủ đề nghiên cứu của nhiều học giả với nhiều công bố khoa học trong các ấn phẩm chuyên khảo và tạp chí khoa học. Một trong những ấn phẩm định kỳ về nhân văn số là tạp chí DSH hay “Digital Scholarship in the Humanities” được ra đời năm 2015 với xếp hạng chỉ số ảnh hưởng là 591 theo danh mục các tạp chí của SCI xuất bản những kết quả nghiên cứu gốc về ứng dụng kỹ thuật số trong các lĩnh vực KHXXNV. Các chủ đề chính của các nghiên cứu được đăng tải trong tạp chí này là những báo cáo nghiên cứu lý thuyết, phương pháp luận, nghiên cứu thực nghiệm và ứng dụng, trong đó trình bày kết quả của các dự án, mô tả và đánh giá các công cụ, kỹ thuật và phương pháp nghiên cứu.

Các giá trị mà các công cụ nhân văn số mang lại cho giới học thuật là rất đa dạng, bao gồm tính mở và khả năng truy cập rộng rãi tới các nguồn lực và nghiên cứu trong lĩnh vực KHXXNV, tăng cường tính hợp tác, tính kết nối và chia sẻ, phát triển tính đa dạng, tính thử nghiệm và khả năng khám phá rộng lớn hơn [Spiro, 2012]. Các công cụ kỹ thuật số cho phép các nhà nghiên cứu xử lý một khối lượng lớn tài liệu một cách

nhANH chóng và toàn diện hơn là xử lý theo cách truyền thống; đồng thời cho phép các nhà nghiên cứu đưa ra các kết luận một cách rõ ràng và minh bạch hơn dựa trên đánh giá các tài liệu và phân tích tổng hợp [Emily & Anne, 2016]. Các lĩnh vực khác nhau có các phương pháp luận cụ thể dựa trên sự ứng dụng của máy tính có thể sẽ tiếp tục tạo ra các công cụ sáng tạo giúp phân tích khối lượng văn bản lớn, tạo cơ hội hợp tác liên ngành trong các lĩnh vực liên quan đến môi trường.

Là một lĩnh vực học thuật giao thoa giữa công nghệ thông tin, kỹ thuật số và các ngành KHxHNV, các chủ đề nghiên cứu nhân văn số rất đa dạng và phụ thuộc vào khuynh hướng phát triển của công nghệ và xã hội, nhiều nhất là về thu thập và quản lý dữ liệu, hình ảnh nhân văn kỹ thuật số [Munster & Terras, 2019], dịch vụ TT - TV phục vụ nghiên cứu nhân văn số, lịch sử kỹ thuật số, văn học kỹ thuật số [Wang, 2018] và bảo tồn di sản văn hóa số [Munster & Terras, 2019; Wang, 2018].

Các phương pháp và công cụ nghiên cứu nhân văn số là một chủ đề nổi bật được nhiều học giả quan tâm, thể hiện mối quan hệ hai chiều, trong đó vừa sử dụng công nghệ để nghiên cứu lại vừa coi ứng dụng công nghệ trong KHxHNV là đối tượng và chủ đề nghiên cứu. Nhân văn số tập trung vào việc sử dụng công nghệ số và phân tích một cách có hệ thống các tài nguyên số trong các lĩnh vực KHxHNV, kiến tạo và sử dụng những cách thức mới trong nghiên cứu học thuật, hợp tác, nghiên cứu liên ngành, giảng dạy và xuất bản ứng dụng công nghệ kỹ thuật số [Drucker, 2013; Terras, 2011]. Các học giả thường quan tâm đến công nghệ ứng dụng trong 5 lĩnh vực nghiên cứu chính của nhân văn số, đó là: (1) các bộ sưu tập kỹ thuật số, lưu trữ và mã hóa văn bản; (2) đọc và phân tích văn bản điện tử; (3) các công nghệ lập bản đồ phân tích không gian địa lý; (4) dữ liệu lớn, điện toán xã hội, tìm

kiếm ý tưởng cộng đồng (croudsourcing) và kết nối mạng; (5) các môi trường trực quan hóa dữ liệu được nhúng trong công nghệ in 3D [Holm, Jarrick & Scott, 2015].

Như vậy, phạm vi, phương pháp và công cụ nghiên cứu trong nhân văn số là rất đa dạng, phong phú và có sự thay đổi cùng với sự phát triển của công nghệ số và những yêu cầu trong nghiên cứu giảng dạy các lĩnh vực KHxHNV.

Hợp tác trong nghiên cứu nhân văn số đang trở nên phổ biến, giúp các nhà nghiên cứu tham gia vào các dự án nghiên cứu số không phụ thuộc vào giới hạn không gian và thời gian. Ba loại hình hợp tác dựa trên quá trình tương tác được đề cập bao gồm: (1) tương tác giữa con người với con người; (2) tương tác giữa con người và các thiết bị, tài liệu và ứng dụng; (3) tương tác giữa máy móc, thiết bị, tài liệu và ứng dụng với nhau [Griffin & Hayler, 2018]. Giữa các nhà nghiên cứu, sự hợp tác liên ngành sẽ giúp mang lại lợi ích to lớn trong nghiên cứu nhân văn số, đặc biệt là giữa các nhà nghiên cứu không có kiến thức chuyên sâu về công nghệ và các chuyên gia công nghệ thông tin. Ví dụ như dự án thiết kế website cho các nghiên cứu dân tộc học thì các học giả, những người bắt buộc phải nghiên cứu phân tích dữ liệu dựa trên những phương pháp mới, hợp tác với những nhân viên phát triển web và phần mềm, những người phải tìm cách chuyển hóa những thông tin tóm tắt và trừu tượng thành các dữ liệu nhị phân, và giữa việc phát triển ứng dụng web dành cho học thuật và dành cho công chúng [Chan & cộng sự, 2017].

Nhìn chung, nhân văn số đã trở thành một trong những chủ đề đang được quan tâm từ nhiều nhà nghiên cứu và những người làm công tác thực tiễn trên thế giới. Sự gia tăng những phát minh và ứng dụng của khoa học công nghệ và kỹ thuật số vào mọi mặt của cuộc sống và công việc của con người chính là động lực để thúc đẩy

sự phát triển của nhân văn số, đòi hỏi các ngành khoa học, xã hội và nhân văn phải phát triển trong bối cảnh mới, dùng những phương pháp mới để nghiên cứu về con người, xã hội và các mối quan hệ liên quan trong kỷ nguyên số. Với khối lượng tài liệu khổng lồ liên quan đến nhân văn số, hướng tiếp cận nghiên cứu toàn diện nhằm khám phá mối quan hệ giữa các bài báo khoa học là rất cần thiết để góp phần xác định các khuynh hướng nghiên cứu liên quan.

2. Nghiên cứu sử dụng trắc lượng thư mục

Trắc lượng thư mục (bibliometrics) hay còn gọi là phân tích thống kê thư mục các công trình đã được công bố. Phương pháp này được nghiên cứu lần đầu tiên bởi Pritchard (1969) và được định nghĩa như là sự ứng dụng của các phương pháp toán học và thống kê nhằm đánh giá định lượng nội dung sách và các tài liệu khác.

Trắc lượng thư mục được sử dụng cho nhiều mục đích, bao gồm giúp nghiên cứu về lịch sử phát triển của các lĩnh vực [Raina & Gupta, 1998] hay điều tra cấu trúc hợp tác của một ngành và liên ngành [Liu & Xia, 2015]. Trắc lượng thư mục thường được dùng để nghiên cứu thống kê dưới nhiều góc độ, bao gồm nghiên cứu phân tích thống kê các đặc điểm của ấn phẩm (tác giả, từ khóa) để định lượng, mô tả và dự đoán quá trình giao tiếp bằng văn bản. Trắc lượng thư mục cũng được dùng để phân tích mạng lưới các yếu tố thư mục liên quan của các tài liệu, bao gồm phân tích đồng tác giả (co-authoring network), đồng trích dẫn (co-citation), đồng xuất hiện của từ khóa (co-occurrence of keywords) và liên kết thư mục (bibliographic coupling). Cụ thể:

- Phân tích đồng tác giả là việc xác định cách thức các nhà nghiên cứu, cơ quan nghiên cứu hoặc quốc gia được kết nối với nhau dựa trên số lượng ấn phẩm mà họ đã đồng xuất bản [Van Eck & Waltman, 2018], giúp hiểu sâu hơn về hoạt động hợp

tác nghiên cứu giữa các tác giả, từ đó đưa ra những định hướng tăng cường hợp tác nghiên cứu.

- Phân tích đồng trích dẫn là sự thể hiện việc hai ấn phẩm được đồng trích dẫn khi có ấn phẩm thứ ba đề cập đồng thời đến chúng [Small, 1973], giúp khám phá sự liên kết giữa các tài liệu, thể hiện khả năng liên quan về mặt ngữ nghĩa của các tài liệu được đồng trích dẫn.

- Phân tích đồng xuất hiện của các từ khóa giúp đánh giá sự xuất hiện của các từ khóa và tần suất mà hai từ khóa được sử dụng đồng thời trong cùng một tài liệu, từ đó gợi ý đến các chủ đề cụ thể được đề cập đến trong nghiên cứu [van Eck et al., 2010].

- Phân tích liên kết thư mục thể hiện việc hai ấn phẩm đồng thời được trích dẫn bởi ấn phẩm thứ ba [Kessler, 1963], cho thấy nếu số lượng trích dẫn/tài liệu tham khảo mà cả hai ấn phẩm cùng sử dụng càng nhiều thì sự liên quan về chủ đề của hai tài liệu này càng cao [van Eck & Waltman, 2018].

Trong lĩnh vực nhân văn số, một số tác giả đã sử dụng phương pháp trắc lượng thư mục để nghiên cứu. Sula (2012) cho rằng, trực quan hóa thông tin thông qua phương pháp phân tích trắc lượng thư mục về nhân văn số thể hiện việc nghiên cứu các mối liên hệ xã hội giữa các học giả trong lĩnh vực nhân văn, giúp tăng cường nhận thức, mở rộng phạm vi thời gian của tài liệu nghiên cứu đã được xuất bản và khám phá dữ liệu tốt hơn. Vì vậy, các phương pháp này nên được sử dụng để có những hiểu biết sâu sắc hơn về tiềm năng, các lĩnh vực nghiên cứu mới để xác định hướng nghiên cứu. Tang, Cheng và Chen (2017) cho rằng, khoa học nhân văn không ngừng mở rộng và phát triển nhưng mức độ tích hợp của các lĩnh vực kiến thức này chưa được khám phá toàn diện. Nhóm tác giả đã tiến hành phân tích thư mục đối với các tài liệu đã xuất bản trong lĩnh vực nhân văn số nhằm kiểm tra mức độ gắn kết trí tuệ theo thời

gian (1989-2014). Kết quả nghiên cứu cho thấy, các ấn phẩm trong lĩnh vực nhân văn số không ngừng phát triển mạnh mẽ, thể hiện tính đa dạng và liên kết liên ngành. Tuy nhiên, mạng lưới đồng tác giả vẫn còn khá rời rạc, với sự hợp tác chủ yếu bị giới hạn bởi ngôn ngữ và ranh giới địa lý.

Việc nghiên cứu tổng quan tài liệu bằng trắc lượng thư mục giúp khám phá kiến thức đã có thuộc một lĩnh vực hoặc liên ngành, xác định các thiếu hụt trong khoa học để đặt ra nhu cầu nghiên cứu bổ sung và đồng thời giúp tích hợp kiến thức, giúp xác định những nghiên cứu mới. Nghiên cứu này sử dụng các phương pháp nghiên cứu trắc lượng thư mục để nghiên cứu các tài liệu xuất bản trong 5 năm gần đây về nhân văn số nói chung sẽ giúp cung cấp một cái nhìn toàn cảnh và cập nhật về lĩnh vực nghiên cứu, giúp các nhà nghiên cứu xác định thế mạnh, và những tồn tại trong nghiên cứu về nhân văn số.

3. Phương pháp nghiên cứu

Phương pháp trắc lượng thư mục được sử dụng trong nghiên cứu này để khảo sát các mối quan hệ giữa các yếu tố thư mục của các bài báo khoa học liên quan đến nhân văn số. Theo đó, chỉ các bài báo tiếng Anh, không giới hạn thời gian xuất bản, từ các tạp chí khoa học có từ khóa “digital humanities” xuất hiện trong mọi trường được tìm kiếm chính xác từ cơ sở dữ liệu Scopus đã được truy xuất.

Kết quả tìm kiếm thu được 1.137 bài báo được trích xuất với các thông tin thư mục để phục vụ cho việc xử lý phân tích dữ liệu. Quá trình xử lý phân tích dữ liệu được tiến hành với những nội dung và công cụ chính sau:

- Lập bản đồ các đặc điểm thư mục của các bài báo đã tìm được, bao gồm thống kê số lượng bài báo, các tác giả, các chủ đề và từ khóa, các bài báo và tạp chí nổi tiếng trong lĩnh vực nhân văn số.

- Sử dụng phần mềm VOSviewer ©, phiên bản 1.6.15 để lập bản đồ các mạng

lưới đồng tác giả, đồng trích dẫn, liên kết thư mục và sự đồng xuất hiện của các từ khóa dựa trên khoảng cách [van Eck & Waltman, 2014]. Mối quan hệ giữa các tác giả, tài liệu hay từ khóa (được thể hiện bằng các nốt tròn trong bản đồ trực quan) của một mạng thư mục được xác định gần đúng bằng khoảng cách giữa chúng, khoảng cách giữa hai nốt càng ngắn và đường kết nối giữa hai nốt càng đậm thì mối quan hệ giữa chúng càng lớn.

- Sử dụng ứng dụng Sciencescape để trích xuất các thông tin mô tả về đặc điểm của các bài báo khoa học

3.1. Kết quả nghiên cứu

- Số lượng bài báo

Hình 2 cho thấy số lượng bài báo về nhân văn số tăng nhanh hàng năm trong hai thập kỷ qua, đặc biệt là từ năm 2017 đến năm 2020. Sự gia tăng này có thể cho thấy khuynh hướng nghiên cứu nhân văn số đến nay và trong những năm tiếp theo; theo đó, dự đoán rằng nhân văn số có thể vẫn là một chủ đề thu hút sự quan tâm của các học giả trong các lĩnh vực khoa học xã hội và nhân văn.



Hình 2. Thống kê số bài báo theo năm xuất bản từ 1999-2021

- Chủ đề nghiên cứu

Các nghiên cứu về nhân văn số tập trung nhiều nhất trong lĩnh vực Khoa học Xã hội, tiếp đến là Nghệ thuật và Khoa học Nhân văn và Khoa học máy tính (Bảng 1). Một

điểm cần chú ý là, các khía cạnh của nhân văn số thường được nghiên cứu trong bối cảnh của các lĩnh vực khoa học nên một tài liệu về nhân văn số có thể được phân loại

và sắp xếp trong một hoặc một số chủ đề của Scopus, dẫn đến tổng số các bài báo trong các chủ đề khác nhau nhiều hơn số lượng bài báo thực tìm được.

Bảng 1. Các lĩnh vực chủ đề của các nghiên cứu về nhân văn số giai đoạn 1999-2021

Chủ đề	Số lượng bài báo
Khoa học xã hội	848
Nghệ thuật và khoa học nhân văn	598
Khoa học máy tính	412
Kỹ thuật	42
Kinh doanh, quản lý và kế toán	21
Các khoa học về trái đất và hành tinh	21
Toán học	19
Khoa học ra quyết định	18
Tâm lý học	17
Khoa học môi trường	16

- Từ khóa của các bài báo khoa học

Để tìm hiểu sâu hơn nội dung nghiên cứu của các bài báo, phân tích thống kê các từ khóa của các bài báo đã được thực hiện

thông qua công cụ Sciencescape. Bảng 2 trình bày 10 từ khóa chính xuất hiện trong các bài báo từ năm 2015 đến năm 2020.

Bảng 2. Tần suất xuất hiện của các từ khóa chính từ năm 2015 đến năm 2020

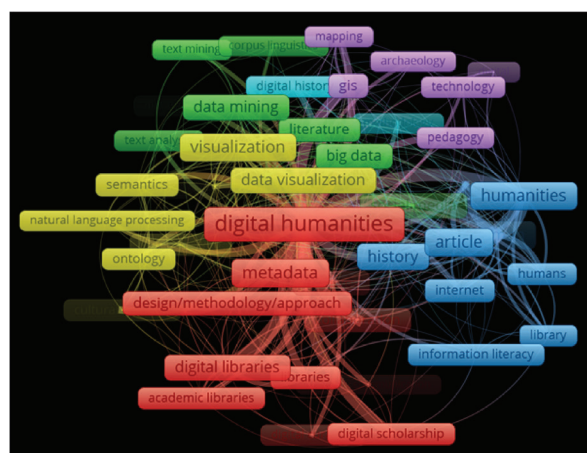
Năm	Từ khóa	Năm	Từ khóa
2020	• digital humanities: 109 bài • crowdsourcing: 4 bài • humanities 4: bài • metadata 4: bài • cultural heritage: 4 bài • digital scholarship: 3 bài • digital history 3: bài • data visualization: 3 bài • distant reading: 3 bài • visualization: 3 bài	2019	• digital humanities: 122 bài • mapping: 8 bài • archives: 6 bài • digital scholarship: 5 bài • data visualization: 5 bài • history: 5 bài • crowdsourcing: 4 bài • visualization: 4 bài • semantic web: 4 bài • text analysis: 4 bài
2018	• digital humanities: 80 bài • collaboration: 8 bài • distant reading: 6 bài • digital scholarship: 4 bài • digital history: 4 bài • crowdsourcing: 4 bài • cultural heritage: 4 bài • digital archive: 4 bài • community: 4 bài • metadata: 3 bài	2017	• digital humanities: 93 bài • pedagogy: 9 bài • collaboration: 6 bài • digital scholarshi: 5 bài • information literacy: 5 bài • digital libraries: 4 bài • gis: 4 bài • embedded librarianship: 4 bài • digital history: 3 bài • archives: 3 bài

2016	• digital humanities: 63 bài • digital scholarship: 4 bài • archives: 4 bài • data visualization: 4 bài • digital libraries: 4 bài • digitization: 4 bài • digital history: 3 bài • academic libraries: 3 bài • text analysis: 3 bài • text mining: 3 bài	2015	• digital humanities: 44 bài • digital libraries: 4 bài • gis: 4 bài • big data: 4 bài • scholarly communication 3 bài • social media: 3 bài • digital scholarship: 2 bài • collaboration: 2 bài • digital history: 2 bài • crowdsourcing: 2 bài
------	--	------	---

Phân tích cho thấy, từ khóa “digital humanities” có khuynh hướng xuất hiện phổ biến nhất ở hầu hết ở các năm với tổng số 654 lần xuất hiện từ giai đoạn 1999-2021, bên cạnh thuật ngữ digital scholarship với 26 lần. Các từ khóa liên quan đến lĩnh vực nghiên cứu ứng dụng của nhân văn số như humanities, human và history cũng xuất hiện nhiều (lần lượt là 47, 39, 27 lần). Các từ khóa được liệt kê cho thấy sự liên quan của nhân văn số với việc sử dụng các công cụ, phương pháp và kỹ thuật số như visualisation, GIS, data visualisation và digital (32, 27, 23, 21 lần xuất hiện), cũng như vai trò của thư viện với nhân văn số thông qua các từ khóa digital libraries (32 lần), metadata (27 lần) và information literacy (10 lần). Như vậy, việc sử dụng các công nghệ số, vai trò của thư viện trong hỗ trợ phát triển nhân văn cho thấy tính ổn định và xu hướng của các hướng nghiên cứu này.

Một trong những phương pháp phân tích cấu trúc của kiến thức khoa học và kỹ thuật là phân tích liên kết đồng xuất hiện của từ khóa [Radhakrishnan, et al., 2017], tập trung vào việc kiểm tra liên kết giữa các từ khóa trong tài liệu. Khi hai hoặc nhiều từ khóa đại diện cho một chủ đề nghiên cứu cụ thể xuất hiện trong cùng một tài liệu thì chúng có mối quan hệ thiết yếu [Chen, et al., 2016]. Sự đồng xuất hiện của hai từ khóa càng nhiều thì mối quan hệ của chúng càng chặt chẽ. Phân tích cho thấy, trong tổng số 3.957 từ khóa, chỉ có 49 từ khóa

xuất hiện 10 lần trở lên (Hình 3). Trong khi đó, có 764 từ khóa xuất hiện hai lần và 373 từ khóa xuất hiện ba lần. Số lượng lớn các từ khóa chỉ được sử dụng một vài lần cho thấy sự thiếu liên tục trong nhiều chủ đề và sự chênh lệch lớn của các trọng tâm nghiên cứu. Các từ khóa nổi bật nhất là digital humanities (trung bình thời gian của năm xuất bản của tài liệu là 2017). Mặc dù tần suất xuất hiện ít hơn, các từ khóa được sử dụng trong thời gian gần đây liên quan đến các công nghệ và kỹ thuật số, như: visualisation, data visualisation, text analysis, mapping, natural language processing, ontology và crowdsourcing đã cho thấy phần nào khuynh hướng ứng dụng công nghệ trong nhân văn số, có thể giúp các nhà nghiên cứu xác định hướng nghiên cứu phù hợp trong tương lai.



Hình 3. Bản đồ sự đồng xuất hiện của các từ khóa

- Tác giả, từ khóa và tạp chí có số bài báo nhiều nhất

Bảng 3 trình bày thống kê 10 tác giả, từ khóa và tạp chí có số lượng bài báo về nhân

văn số nhiều nhất. Các tác giả nổi bật có số lượng bài viết nhiều nhất về nhân văn số là Poole và Terras (6 bài), Blanke, Gregory, Risam và Warwick (5 bài).

Bảng 3. Thống kê số lượng bài báo cho các tác giả, từ khóa và tạp chí phổ biến nhất

Tác giả nổi bật	Từ khóa nổi bật	Tạp chí nổi bật
• Poole A.A. (6) • Terras M. (6) • Blanke T. (5) • Gregory I.N. (5) • Risam R. (5) • Warwick C. (5) • Chen C.-M. (4) • Ciula A. (4) • Edmond J. (4) • Garwood D.A. (4)	• digital humanities (654) • digital scholarship (26) • collaboration (25) • digital history (21) • archives (19) • data visualization (19) • digital libraries (19) • pedagogy (19) • crowdsourcing (18) • humanities (18)	• digital scholarship in the humanities (55) • digital humanities quarterly (36) • college and undergraduate libraries (30) • literary and linguistic computing (24) • journal of documentation (18) • proceedings of the association for information science and technology (15) • library trends (14) • historical social research (13) • digital studies/ le champ numerique (12) • digital library perspectives (11)

Các từ khóa liên quan đến công cụ và phương pháp nghiên cứu trao đổi học thuật cũng được dùng đa dạng như: data visualization, crowdsourcing, mapping, semantic web, text analysis và text mining. Các từ khóa liên quan đến thư viện như: digital libraries, metadata, academic libraries và digitization cho thấy vai trò và sự tham gia tích cực của thư viện trong nghiên cứu và hỗ trợ phát triển nhân văn số trong giáo dục bậc đại học. Tần suất xuất hiện của từ khóa cũng cho thấy những lĩnh vực nghiên cứu được quan tâm, bao gồm pedagogy, humanities, cultural heritage và history.

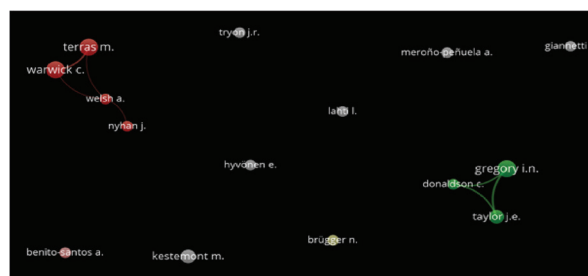
Tạp chí Digital Scholarship in the Humanities và Digital humanities quarterly được xem như những tạp chí chuyên về nhân văn số phổ biến nhất. Bên cạnh đó, có thể thấy có nhiều tạp chí liên quan đến lĩnh vực Khoa học Thông tin và Thư viện, cho thấy vai trò của hoạt động thư viện trong nghiên cứu và phát triển nhân văn số hiện nay.

- Quan hệ đồng tác giả giữa các tác giả cá nhân và quốc gia

Để tìm hiểu khuynh hướng hợp tác trong nghiên cứu nhân văn số, nghiên cứu này đã tiến hành phân tích quan hệ đồng tác giả giữa các tác giả cá nhân và giữa các

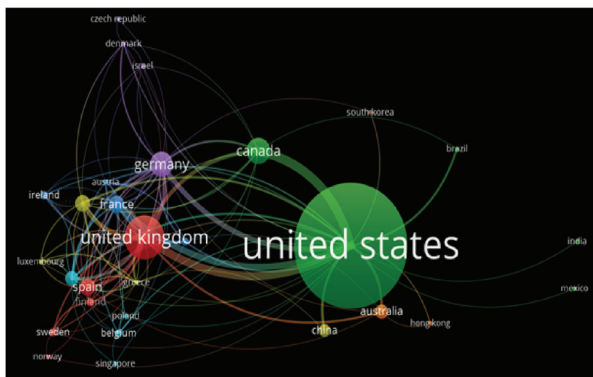
quốc gia. Theo đó, kết quả phân tích giúp nâng cao hiểu biết về sự hợp tác nghiên cứu, đồng thời giúp phát hiện ra những nhà nghiên cứu có nhiều ảnh hưởng [Wang, 2018]. Hình 4 trình bày bản đồ mạng lưới đồng tác giả, trong đó các tác giả được thể hiện dưới dạng các nốt với ít nhất ba bài báo, nhãn nốt là họ của tác giả, kích thước nốt cho biết số lượng bài báo đã xuất bản. Liên kết giữa hai nốt đại diện cho mối quan hệ hợp tác giữa hai tác giả, và độ đậm của đường liên kết tượng trưng cho cường độ hợp tác.

Trong tổng số 2.216 tác giả, có 52 tác giả có tối thiểu 3 bài báo, có 8 nhóm đồng tác giả tối thiểu 2 bài báo. Các nhóm có mối quan hệ hợp tác chặt chẽ nhất là nhóm của các tác giả (1) Terras, Warwick, Welsh và Nyhan và (2) Donaldson, Gregory và Taylor.



Hình 4. Sự liên kết của các tác giả cá nhân (n=70)

Bản đồ trực quan hóa mạng lưới hợp tác giữa các quốc gia (Hình 5) cho thấy, trong tổng số 72 quốc gia, có 33 quốc gia là đồng tác giả quốc tế với điều kiện mỗi nước có ít nhất là 5 tài liệu. Trong bản đồ, mỗi quốc gia được thể hiện bằng một nhãn tên nước và một nốt tròn, kích thước của các nốt tròn thể hiện số lượng bài báo của tác giả thuộc nước đó, quốc gia càng quan trọng thì nốt tròn và nhãn càng lớn. Độ dày của các đường nối các quốc gia thể hiện quy mô của sự cộng tác, ví dụ quy mô hợp tác mạnh nhất là giữa Hoa Kỳ và Canada với độ mạnh liên kết (link strength) là 16 và tiếp đến là Hoa Kỳ và Anh với độ mạnh liên kết là 14. Hầu hết các kết quả nghiên cứu đều đến từ Hoa Kỳ với 441 bài, Anh có 139 bài, Canada có 78 bài và Đức có 74 bài. Trong tổng số 8 nhóm đồng tác giả, các nước có quan hệ mạng lưới hợp tác nổi bật nhất thuộc về nhóm 1 màu đỏ (gồm: Anh, Phần Lan, Na Uy, Tây Ban Nha và Thụy Điển), và nhóm 2 màu xanh lá cây (gồm: Hoa Kỳ, Mexico, Ấn Độ, Canada và Brazil) với mỗi nhóm có 5 bài báo. Một số nước châu Á tham gia tích cực vào mạng lưới hợp tác là: Trung Quốc, Ấn Độ, Singapore và Hàn Quốc; tuy nhiên Việt Nam chưa xuất hiện trên bản đồ hợp tác nghiên cứu về nhân văn số này. Nhìn chung, kết quả phân tích cho thấy mạng lưới hợp tác giữa các quốc gia còn khá rời rạc và có khuynh hướng bị giới hạn bởi ranh giới địa lý và trình độ phát triển của mỗi quốc gia.



Hình 5. Mạng lưới hợp tác giữa các quốc gia

Kết luận

Nghiên cứu này đã sử dụng phương pháp trắc lượng thư mục với sự trợ giúp của một số ứng dụng giúp thống kê và trực quan hóa dữ liệu nhằm khám phá các khuynh hướng nghiên cứu trong nhân văn số. Kết quả phân tích cho thấy, về nội dung nghiên cứu, nhân văn số là khuynh hướng nghiên cứu được phát triển liên tục, thể hiện tính đa dạng và gắn kết liên ngành mạnh mẽ. Các hướng nghiên cứu của nhân văn số ngày càng mang tính chuyên sâu cao, tập trung vào các lĩnh vực như lịch sử, khảo cổ học, ngôn ngữ, sử phạm, di sản văn hóa số và các khoa học nhân văn. Việc sử dụng các công cụ, phương pháp và kỹ thuật số trong nghiên cứu ngày càng đa dạng và phổ biến. Vai trò của thư viện trong hỗ trợ nhân văn số ngày càng được khẳng định thể hiện qua các bài báo và chủ đề liên quan đến hoạt động hỗ trợ nghiên cứu của thư viện. Kết quả phân tích cũng xác định các tác giả và tạp chí nổi bật, giúp các nhà khoa học tìm được những tham khảo quan trọng cho quá trình nghiên cứu và xuất bản.

Kết quả nghiên cứu cũng cho thấy, khuynh hướng trong hợp tác nghiên cứu, giữa các tác giả, cũng như giữa các quốc gia với nhau. Mặc dù các tác giả từ các nước đang phát triển đã có nhiều công trình nghiên cứu liên quan, vai trò nổi bật nhất vẫn đến từ các tác giả ở các quốc gia phát triển. Việt Nam chưa có trong bản đồ hợp tác nói trên, nên đây là lĩnh vực nghiên cứu tiềm năng cho các nhà khoa học trong nghiên cứu và xuất bản nhân văn số với những chủ đề phù hợp với xu hướng của khu vực và thế giới.

Kết quả nghiên cứu được trình bày trong bài báo này phần nào giúp các nhà khoa học có một nhìn nhận tổng quát về tình hình nghiên cứu nhân văn số và giúp định hướng tìm kiếm những chủ đề nghiên cứu phù hợp, xác định các phương pháp và công cụ nghiên cứu cũng như tìm kiếm các hỗ trợ từ các cá nhân và tổ chức liên quan.

TÀI LIỆU THAM KHẢO

1. Busa, R. (1980). The Annals of Humanities Computing: The Index Thomisticus, Computers and the Humanities, 14, 83–90.
2. Chan, A.; Chenhall, R.; Kohn, T. & Stevens, C. (2017). Interdisciplinary Collaboration and Brokerage in the Digital Humanities. DHQ: Digital Humanities Quarterly. 11. Truy cập từ <http://www.digitalhumanities.org/dhq/vol/11/3/000336/000336.html>
3. Drucker, J. (2013). Intro to Digital Humanities: Introduction. UCLA Center for Digital Humanities. Truy cập từ http://dh101.humanities.ucla.edu/?page_id=13.
4. Emily, G. & Anne, S. (2016). Benefits and applications of interdisciplinary digital tools for environmental meta-reviews and analyses. Environmental Research Letters, 11 (9), 1-13. Truy cập từ <https://iopscience.iop.org/article/10.1088/1748-9326/11/9/093001/pdf>
5. Gavin, M. & Smith, K. M. (2012). An Interview with Brett Bobley. Gold, M.K. (Ed.), Debates in the Digital Humanities (61-66). London: University of Minnesota Press.
6. Griffin, G. & Hayler, M. S. (2018). Collaboration in digital humanities research - Persisting silences. Digital Humanities Quarterly, 12(1). Truy cập từ <http://www.digitalhumanities.org/dhq/>
7. Hockey, S. (2004). The History of Humanities Computing. Schreibman, S., Siemens, R. & Unsworth, J. (Ed.), A Companion to Digital Humanities (3-19). Victoria: Blackwell Publishing.
8. Holm, P., Jarrick, A. & Scott, D. (2015). Humanities world report 2015. Basingstoke, England; New York: Palgrave Macmillan.
9. Kessler MM. "Bibliographie coupling between scientific papers". American Documentation, 1963; 14(1): 10-11.
10. Liu, P., & Xia, H. (2015). Structure and evolution of co-authorship network in an interdisciplinary research field. Scientometrics, 103(1), 101–134.
11. Münster, S. & Terras, M. (2019). The visual side of digital humanities: a survey on topics, researchers, and epistemic cultures. Digital Scholarship in the Humanities, 0(0), 1-24. Truy cập từ <https://doi.org/10.1093/llc/fqz022>
12. Pritchard, A. (1969). Statistical bibliography or bibliometrics?. Journal of Documentation, 25(4), 348-349.
13. Raina, D., & Gupta, B. (1998). Four aspects of the institutionalization of physics research in India (1990–1950): Substantiating the claims of historical sociology through bibliometrics. Scientometrics, 42(1), 17-40.
14. Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2008). A companion to digital humanities. John Wiley & Sons.
15. Small H. "Co-citation in the scientific literature: A new measure of the relationship between two documents". Journal of the American Society for Information Science, 1973; 24(3): 265-269.
16. Spiro, L. (2012). "This Is Why We Fight": Defining the Values of the Digital Humanities (pp. 16-34). University of Minnesota.
17. Sula, C. A. (2012). Visualizing social connections in the humanities: Beyond bibliometrics. Bulletin of the American Society for Information Science and Technology, 38(4), 31-35.
18. Van Eck N.J., Waltman L., 2018, Manual for VOSviewer version 1.6.8, CWTS Meaningful Metrics, Universiteit Leiden
19. Van Eck, N. J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. Journal of the American Society for Information Science and Technology, 61(12), 2405-2416.
20. Tang, M. C., Cheng, Y. J., & Chen, K. H. (2017). A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses. Scientometrics, 113(2), 985-1008.
21. Wang, Q. (2018), "Distribution features and intellectual structures of digital humanities: A bibliometric analysis", Journal of Documentation, 74(1), 223-246

(Ngày Tòa soạn nhận được bài: 15-01-2021;
Ngày phản biện đánh giá: 16-3-2021; Ngày
chấp nhận đăng: 15-5-2021).